

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/033,129C
Source: IFW/6
Date Processed by STIC: 8/15/05

ENTERED

BEST AVAILABLE COPY

**CRF ERRORS Edited by the STIC Systems
Branch**

Serial Number: 10/033,129C

CRF Edit Date: 8/16/05

Edited by: JW

Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

Corrected the SEQ ID NO. Sequence numbers edited were:

Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

Inserted mandatory headings/numeric identifiers, specifically:

Moved responses to same line as heading/numeric identifier, specifically:

Other:

Sequence 17 - corrected amino acid numbering



IFW16

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/033,129C

DATE: 08/16/2005

TIME: 11:31:09

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08162005\J033129C.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: CHOO, Yen

6 KLUG, Aaron

7 SANCHEZ GARCIA, Isidro

9 (ii) TITLE OF INVENTION: Improvements in or Relating to
10 Binding Proteins for Recognition of DNA

12 (iii) NUMBER OF SEQUENCES: 125

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.

16 (B) STREET: 1100 New York Avenue, N.W.

17 (C) CITY: Washington

18 (D) STATE: D.C.

19 (E) COUNTRY: USA

20 (F) ZIP: 20005-3918

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette

24 (B) COMPUTER: IBM PC compatible

25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

26 (D) SOFTWARE: Word Perfect

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/10/033,129C

C--> 30 (B) FILING DATE: 27-Dec-2001

31 (C) CLASSIFICATION:

48 (vii) PRIOR APPLICATION DATA:

W--> 34 (A) APPLICATION NUMBER: US 08/793,408

35 (B) FILING DATE: 02-JUN-1997

W--> 37 (A) APPLICATION NUMBER: PCT/GB95/01949

38 (B) FILING DATE: 17-AUG-1995

W--> 41 (A) APPLICATION NUMBER: GB 9514698.1

42 (B) FILING DATE: 18-JUL-1995

W--> 45 (A) APPLICATION NUMBER: GB 9422534.9

46 (B) FILING DATE: 08-NOV-1994

W--> 49 (A) APPLICATION NUMBER: GB 9416880.4

50 (B) FILING DATE: 20-AUG-1994

52 (2) INFORMATION FOR SEQ ID NO: 1:

54 (i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 60 base pairs

56 (B) TYPE: nucleic acid

57 (C) STRANDEDNESS: single

58 (D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

65 CTCCCTGCAGT TGGACCTGTG CCATGGCCGG CTGGGCCGCA TAGAATGGAA

50

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/033,129C

DATE: 08/16/2005

TIME: 11:31:09

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08162005\J033129C.raw

66 CAACTAAAGC 60
 68 (2) INFORMATION FOR SEQ ID NO: 2:
 70 (i) SEQUENCE CHARACTERISTICS:
C--> 71 (A) LENGTH: 92 amino acids
 72 (B) TYPE: amino acid
 73 (C) STRANDEDNESS:
 74 (D) TOPOLOGY: unknown
 76 (ii) MOLECULE TYPE: protein
 78 (ix) FEATURE:
 79 (A) NAME/KEY: variable amino acid residue
 80 (B) LOCATION: 57
 81 (C) IDENTIFICATION METHOD: by experiment
 82 (D) OTHER INFORMATION: Arg or Lys
 84 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 86 Met Ala Glu Glu Arg Pro Tyr Ala Cys Pro
 87 5 10
 88 Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
 89 15 20
 90 Ser Asp Glu Leu Thr Arg His Ile Arg Ile
 91 25 30
 92 His Thr Gly Gln Lys Pro Phe Gln Cys Arg
 93 35 40
W--> 94 Ile Cys Met Arg Asn Phe Ser Xaa Xaa Xaa
 95 45 50
 96 Xaa Leu Xaa Xaa His Xaa Xaa Thr His Thr
 97 55 60
 98 Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys
 99 65 70
 100 Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg
 101 75 80
 102 Lys Arg His Thr Lys Ile His Leu Arg Gln
 103 85 90
 104 Lys Asp
 106 (2) INFORMATION FOR SEQ ID NO: 3:
 108 (i) SEQUENCE CHARACTERISTICS:
 109 (A) LENGTH: 26 base pairs
 110 (B) TYPE: nucleic acid
 111 (C) STRANDEDNESS: single
 112 (D) TOPOLOGY: linear
 114 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 116 TATGACTTG ATGGGAGACC GCCTGG 26
 118 (2) INFORMATION FOR SEQ ID NO: 4:
 120 (i) SEQUENCE CHARACTERISTICS:
 121 (A) LENGTH: 28 base pairs
 122 (B) TYPE: nucleic acid
 123 (C) STRANDEDNESS: single
 124 (D) TOPOLOGY: linear
 126 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 128 AATTCCAGGC GGTCTCCCAT CCAAGTCA 28

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/033,129C DATE: 08/16/2005
 TIME: 11:31:09

Input Set : A:\PTO.AMC.txt
 Output Set: N:\CRF4\08162005\J033129C.raw

```

130 (2) INFORMATION FOR SEQ ID NO: 5:
132     (i) SEQUENCE CHARACTERISTICS:
133         (A) LENGTH: 21 base pairs
134         (B) TYPE: nucleic acid
135         (C) STRANDEDNESS: single
136         (D) TOPOLOGY: linear
138     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
140 TATATAGCGT GGGCGTATAT A                                21
142 (2) INFORMATION FOR SEQ ID NO: 6:
144     (i) SEQUENCE CHARACTERISTICS:
145         (A) LENGTH: 24 base pairs
146         (B) TYPE: nucleic acid
147         (C) STRANDEDNESS: single
148         (D) TOPOLOGY: linear
150     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
152 GCGTATATAAC GCCCACGCTA TATA                            24
154 (2) INFORMATION FOR SEQ ID NO: 7:
156     (i) SEQUENCE CHARACTERISTICS:
157         (A) LENGTH: 21 base pairs
158         (B) TYPE: nucleic acid
159         (C) STRANDEDNESS: single
160         (D) TOPOLOGY: linear
162     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
164 TATATAGCGN NNGCGTATAT A                                21
166 (2) INFORMATION FOR SEQ ID NO: 8:
168     (i) SEQUENCE CHARACTERISTICS:
169         (A) LENGTH: 24 base pairs
170         (B) TYPE: nucleic acid
171         (C) STRANDEDNESS: single
172         (D) TOPOLOGY: linear
174     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
176 GCGTATATAAC GCNNNCGCTA TATA                            24
179 (2) INFORMATION FOR SEQ ID NO: 9:
181     (i) SEQUENCE CHARACTERISTICS:
182         (A) LENGTH: 33 base pairs
183         (B) TYPE: nucleic acid
184         (C) STRANDEDNESS: single
185         (D) TOPOLOGY: linear
187     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
189 TTCCATGGAG ACGCAGAAC CCTTCAGCGG CCA                  33
191 (2) INFORMATION FOR SEQ ID NO: 10:
193     (i) SEQUENCE CHARACTERISTICS:
194         (A) LENGTH: 33 base pairs
195         (B) TYPE: nucleic acid
196         (C) STRANDEDNESS: single
197         (D) TOPOLOGY: linear
199     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
201 TTCCATGGAG ACGCAGGTGA GTTCCTCACG CCA                  33
203 (2) INFORMATION FOR SEQ ID NO: 11:

```

RAW SEQUENCE LISTING DATE: 08/16/2005
PATENT APPLICATION: US/10/033,129C TIME: 11:31:09

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\08162005\J033129C.raw

205 (i) SEQUENCE CHARACTERISTICS:
206 (A) LENGTH: 33 base pairs
207 (B) TYPE: nucleic acid
208 (C) STRANDEDNESS: single
209 (D) TOPOLOGY: linear

211 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
213 CCCCTTTCTC TTCCAGAAGC CCTTCAGCGG CCA

33

215 (2) INFORMATION FOR SEQ ID NO: 12:
216 (i) SEQUENCE CHARACTERISTICS:
217 (A) LENGTH: 33 amino acids
218 (B) TYPE: amino acid
219 (C) STRANDEDNESS:
220 (D) TOPOLOGY: unknown

221 (D) TOPOLOGY: unknown
223 (ii) MOLECULE TYPE: peptide
225 (vi) SEQUENCE DESCRIPTION: SEQ ID NO: 12.

225 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg
226 5

228 5 10
 230 Ile Cys Met Arg Asn Phe Ser Asp Arg Ser
 231 15 20
 232 Ser Leu Thr Arg His Thr Arg His Thr Gly

233 Ser Leu Thr Arg His Thr Arg His Thr Gly
234 25 30
235 Gly Lys Pro

236 Glu Lys Pro

239 (2) INFORMATION FOR SEQ ID NO: 13:
341 (i) SEQUENCE CHARACTERISTICS:

241 (i) SEQUENCE CHARACTERISTICS:
242 (A) LENGTH: 33 amino acids
243 (B) TYPE: amino acid
244 (C) STRANDEDNESS:
245 (D) TOPOLOGY: unknown
247 (ii) MOLECULE TYPE: peptide

247 (ii) MOLECULE TYPE: peptide
249 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
251 Met Ala Glu Glu Ile Pro Phe Gln Cys Arg

251 Met Ala Glu Glu Lys Pro Phe Gln Cys Arg
 252 5 10
 253 Glu Ser Asp Asn Ile Phe Ser Gly

260 Glu Lys Pro
260 (S) INFORMATION FOR SEQ ID NO: 14

262 (2) INFORMA

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH OF amino acid

(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

272 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

274 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
275 5 10

277 Ser Gln Gly Gly Asn Leu Val Arg His Leu
278 15 30

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/033,129C

DATE: 08/16/2005

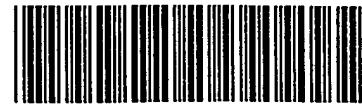
TIME: 11:31:09

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08162005\J033129C.raw

280 Arg His Thr Gly Glu Lys Pro
281 25
283 (2) INFORMATION FOR SEQ ID NO: 15:
285 (i) SEQUENCE CHARACTERISTICS:
286 (A) LENGTH: 26 amino acids
287 (B) TYPE: amino acid
288 (C) STRANDEDNESS:
289 (D) TOPOLOGY: unknown
291 (ii) MOLECULE TYPE: peptide
293 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
295 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
296 5 10
299 Ser Gln Ala Gln Thr Leu Gln Arg His Leu
300 15 20
302 Lys His Thr Gly Glu Lys
303 25
305 (2) INFORMATION FOR SEQ ID NO: 16:
307 (i) SEQUENCE CHARACTERISTICS:
308 (A) LENGTH: 26 amino acids
309 (B) TYPE: amino acid
310 (C) STRANDEDNESS:
311 (D) TOPOLOGY: unknown
313 (ii) MOLECULE TYPE: peptide
315 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
317 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
318 5 10
320 Ser Gln Ala Ala Thr Leu Gln Arg His Leu
321 15 20
323 Lys His Thr Gly Glu Lys
324 25
326 (2) INFORMATION FOR SEQ ID NO: 17:
328 (i) SEQUENCE CHARACTERISTICS:
329 (A) LENGTH: 26 amino acids
330 (B) TYPE: amino acid
331 (C) STRANDEDNESS:
332 (D) TOPOLOGY: unknown
334 (ii) MOLECULE TYPE: peptide
336 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
338 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
339 5 10
341 Ser Gln Ala Gln Asp Leu Gln Arg His Leu
342 15 20
344 Lys His Thr Gly Glu Lys
345 25
347 (2) INFORMATION FOR SEQ ID NO: 18:
349 (i) SEQUENCE CHARACTERISTICS:
350 (A) LENGTH: 89 amino acids
351 (B) TYPE: amino acid
352 (C) STRANDEDNESS:

**Raw Sequence Listing before editing,
for reference only**



IFW16

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/033,129C

DATE: 08/15/2005
 TIME: 08:35:39

Input Set : A:\PTO.DA.txt
 Output Set: N:\CRF4\08152005\J033129C.raw

SEQUENCE LISTING

Cor...
...order

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: CHOO, Yen
 6 KLUG, Aaron

7 SANCHEZ GARCIA, Isidro

9 (ii) TITLE OF INVENTION: Improvements in or Relating to
 10 Binding Proteins for Recognition of DNA

12 (iii) NUMBER OF SEQUENCES: 125

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
 16 (B) STREET: 1100 New York Avenue, N.W.

17 (C) CITY: Washington

18 (D) STATE: D.C.

19 (E) COUNTRY: USA

20 (F) ZIP: 20005-3918

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette

24 (B) COMPUTER: IBM PC compatible

25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

26 (D) SOFTWARE: Word Perfect

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/10/033,129C

C--> 30 (B) FILING DATE: 27-Dec-2001

31 (C) CLASSIFICATION:

45 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: PCT/GB95/01949

35 (B) FILING DATE: 17-AUG-1995

38 (A) APPLICATION NUMBER: GB 9514698.1

39 (B) FILING DATE: 18-JUL-1995

42 (A) APPLICATION NUMBER: GB 9422534.9

43 (B) FILING DATE: 08-NOV-1994

46 (A) APPLICATION NUMBER: GB 9416880.4

47 (B) FILING DATE: 20-AUG-1994

Does Not Comply
Corrected Diskette Neede

Does Not Comply
Corrected Diskette Neede

ERRORED SEQUENCES

323 (2) INFORMATION FOR SEQ ID NO: 17:

325 (i) SEQUENCE CHARACTERISTICS:

326 (A) LENGTH: 26 amino acids

327 (B) TYPE: amino acid

328 (C) STRANDEDNESS:

329 (D) TOPOLOGY: unknown

β.2

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/033,129C

DATE: 08/15/2005
TIME: 08:35:39

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\08152005\J033129C.raw

331 (ii) MOLECULE TYPE: peptide
333 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
335 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
336 5 10
338 Ser Gln Ala Gln Asp Leu Gln Arg His Leu
339 15 20
341 Lys His Thr Gly Glu Lys
E--> 342 25 25

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/033,129C

DATE: 08/15/2005

TIME: 08:35:41

Input Set : A:\PTO.DA.txt

Output Set: N:\CRF4\08152005\J033129C.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:68 M:220 C: Keyword misspelled or invalid format, [(A) LENGTH:]
L:91 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:40
M:341 Repeated in SeqNo=2
L:342 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:17

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS

IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

FADED TEXT OR DRAWING

BLURRED OR ILLEGIBLE TEXT OR DRAWING

SKEWED/SLANTED IMAGES

COLOR OR BLACK AND WHITE PHOTOGRAPHS

GRAY SCALE DOCUMENTS

LINES OR MARKS ON ORIGINAL DOCUMENT

REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.